

What is claimed is:

1 1. A method for a wide-area file system, including a
2 plurality of replicas for a file, wherein each replica
3 of the file and parent directories for the file are at
4 each of a plurality of nodes, the method comprising:
5 propagating an update to a replica of the file to
6 other replicas of the file; and
7 in response receiving a propagated update to a
8 replica at a node, updating the replica for the file
9 at the node.

1 2. The method according to claim 1, wherein each
2 replica has a backpointer including an identification
3 of a parent directory for the file and a name of the
4 file in the parent directory.

1 3. The method according to claim 2, wherein the
2 parent directories are modified when the backpointer
3 for a replica at a node is not consistent with the
4 parent directories for the replica at the node.

1 4. The method according to claim 3, wherein
2 modifying the parent directories occurs only after a
3 delay.

1 5. The method according to claim 3, wherein multiple
2 modifications to the parent directories at the node
3 are performed according to an order in which
4 corresponding updates occur.

1 6. The method according to claim 3, wherein a
2 modification is performed at the node and an earlier
3 inconsistent modification is ignored.

1 7. The method according to claim 2, wherein a
2 directory operation affects the backpointer for the
3 file.

1 8. The method according to claim 7, wherein the
2 directory operation is selected from a group
3 consisting of rename, link and unlink.

1 9. The method according to claim 8, wherein when the
2 backpointer for a replica at a node is not consistent
3 with the parent directories for the replica at the
4 node, further comprising modifying the parent
5 directories to be consistent with the backpointer.

1 10. The method according to claim 1, wherein the
2 replicas include replicas of a first type and of a
3 second type wherein locations of replicas of the first
4 type are registered in a parent directory for a file.

1 11. The method according to claim 1, wherein the
2 replicas include replicas of a first type and of a
3 second type and wherein in response to a user
4 accessing a file at a node, the method further
5 comprises steps of replicating the parent directory
6 for the file at the node and forming of replica of the
7 second type at the node.

1 12. The method according to claim 11, wherein a
2 minimum number of replicas of the first type are
3 maintained according to a minimum replication factor
4 for the corresponding file.

1 13. The method according to claim 1, wherein a
2 replica is deleted by marking the replica as invalid.

1 14. The method according to claim 13, wherein said
2 marking the replica as invalid comprises removing the
3 backpointer for the replica.

1 15. The method according to claim 13, further
2 comprising periodically removing replicas marked as
3 invalid.

1 16. A method for a wide-area file system, including a
2 plurality of replicas for a file, and wherein upon
3 access of the file by a user at a node, the method
4 comprises steps of:

5 replicating parent directories for the file at
6 the node; and

7 forming a replica of the file at the node wherein
8 the replica includes a backpointer having an
9 identification of a parent directory for the file and
10 a name of the file within the parent directory.

1 17. The method according to claim 16, further
2 comprising propagating an update to a replica of the
3 file to other replicas of the file.

1 18. The method according to claim 17, wherein the
2 update is forwarded according to a graph for the file.

1 19. The method according to claim 17, wherein in
2 response receiving a propagated update to a replica at
3 a node, the node updates parent directories for the
4 file at the node.

1 20. The method according to claim 17, wherein when a
2 backpointer for a replica at a node is not consistent
3 with parent directories for the replica at the node,
4 the method further comprises modifying the parent
5 directories to be consistent with the backpointer.

1 21. The method according to claim 20, wherein said
2 modifying is performed only after a delay.

1 22. The method according to claim 20, wherein
2 multiple modifications to the parent directories at
3 the node are performed according to an order in which
4 corresponding updates occur.

1 23. The method according to claim 20, wherein a
2 modification is performed at the node and an earlier
3 inconsistent modification is ignored.

1 24. The method according to claim 17, wherein a
2 directory operation affects the backpointer for the
3 file.

1 25. The method according to claim 24, wherein the
2 directory operation is selected from a group
3 consisting of rename, link and unlink.

1 26. The method according to claim 24, wherein when
2 the backpointer for a replica at a node is not
3 consistent with parent directories for the replica at
4 the node, the method further comprises modifying the
5 parent directories to be consistent with the
6 backpointer.

1 27. The method according to claim 16, wherein a
2 minimum number of replicas for a file are maintained
3 according to a minimum replication factor for the
4 file.

1 28. The method according to claim 27, wherein links
2 between replicas of a file form a graph for the file
3 and wherein updates to the file are propagated along
4 the graph.

1 29. The method according to claim 16, wherein a
2 replica is deleted by marking the replica as invalid.

1 30. The method according to claim 29, wherein said
2 marking the replica as invalid comprises removing the
3 backpointer for the replica.

1 31. The method according to claim 29, further
2 comprising periodically removing replicas marked as
3 invalid.

1 32. The method according to claim 16, further
2 comprising locating a parent directory of the file.

1 33. The method according to claim 32, wherein said
2 replicating the parent directories comprises copying
3 contents of the parent directory located by said
4 locating.

1 34. The method according to claim 33, wherein said
2 locating and copying are performed recursively,
3 thereby traversing a pathname for the file.

1 35. The method according to claim 33, wherein said
2 forming the replica of the file comprises a copying
3 contents of a replica of the file located by said
4 locating.

1 36. A method for a wide-area file system including a
2 first type of file replica and a second type of file
3 replica wherein locations of replicas of the first
4 type are registered in a parent directory for a file,
5 and wherein upon access of a file by a user at a node,
6 the method comprises steps of:

7 replicating the parent directory for the file at
8 the node; and

9 forming a replica of the second type at the node
10 wherein the replica of the second type includes a
11 backpointer having an identification of the parent
12 directory for the file and a name of the file within
13 the parent directory.

1 37. The method according to claim 36, further
2 comprising propagating an update to a replica of the
3 file to other replicas of the file.

1 38. The method according to claim 37, wherein the
2 update is forwarded according to a graph for the file.

1 39. The method according to claim 37, wherein in
2 response receiving a propagated update to a replica at
3 a node, the node updates parent directories for the
4 file at the node.

1 40. The method according to claim 37, wherein when a
2 backpointer for a replica at a node is not consistent
3 with parent directories for the replica at the node,
4 the method further comprises modifying the parent
5 directories to be consistent with the backpointer.

1 41. The method according to claim 40, wherein said
2 modifying is performed only after a delay.

1 42. The method according to claim 40, wherein
2 multiple modifications to the parent directories at
3 the node are performed according to an order in which
4 corresponding updates occur.

1 43. The method according to claim 40, wherein a
2 modification is performed at the node and an earlier
3 inconsistent modification is ignored.

1 44. The method according to claim 37, wherein a
2 directory operation is performed on the backpointer
3 for the file.

1 45. The method according to claim 44, wherein the
2 directory operation is selected from the group
3 consisting of rename, link and unlink.

1 46. The method according to claim 44, wherein when
2 the backpointer for a replica at a node is not
3 consistent with parent directories for the replica at
4 the node, the method further comprises modifying the
5 parent directories to be consistent with the
6 backpointer.

1 47. The method according to claim 36, wherein a
2 minimum number of replicas of the first type are
3 maintained according to a minimum replication factor
4 for the corresponding file.

1 48. The method according to claim 47, wherein links
2 between replicas of a file form a graph for the file
3 and wherein updates to the file are propagated along
4 the graph.

1 49. The method according to claim 36, wherein a
2 replica is deleted by marking the replica as invalid.

1 50. The method according to claim 49, wherein said
2 marking the replica as invalid comprises removing the
3 backpointer for the replica.

1 51. The method according to claim 49, further
2 comprising periodically removing replicas marked as
3 invalid.

1 52. The method according to claim 36, further
2 comprising locating a parent directory of the file.

1 53. The method according to claim 52, wherein said
2 replicating the parent directories comprises copying
3 contents of the parent directory located by said
4 locating.

1 54. The method according to claim 53, wherein said
2 locating and copying are performed recursively,
3 thereby traversing a pathname for the file.

1 55. The method according to claim 53, wherein said
2 forming the replica of the file comprises a copying
3 contents of a replica of the file located by said
4 locating.

1 56. A system including a plurality of nodes for
2 storing replicas of files wherein for each replica at
3 a node, the node stores parent directories for the
4 file and a backpointer having an identification of a
5 parent directory for the file and wherein updates to
6 replicas of the file are propagated to other replicas
7 of the file.

1 57. The system according to claim 56, wherein in
2 response to receiving a propagated update to a replica

3 at a node, the node updates the parent directories for
4 the file at the node.

1 58. The system according to claim 56, wherein when a
2 backpointer for a replica at a node is not consistent
3 with the parent directories for the replica at the
4 node, the node modifies the parent directories to be
5 consistent with the backpointer.

1 59. The system according to claim 58, wherein the
2 node modifies the parent directories to be consistent
3 with the backpointer only after a delay.

1 60. The method according to claim 58, wherein
2 multiple modifications to the parent directories at
3 the node are performed according to an order in which
4 corresponding updates occur.

1 61. The method according to claim 58, wherein a
2 modification is performed at the node and an earlier
3 inconsistent modification is ignored.

1 62. The method according to claim 56, wherein a
2 directory operation is affects the backpointer for the
3 file.

1 63. The method according to claim 62, wherein the
2 directory operation is selected from a group
3 consisting of rename, link and unlink.

1 64. The method according to claim 62, wherein when a
2 backpointer for a replica at a node is not consistent

3 with the parent directories for the replica at the
4 node, the node modifies the parent directories to be
5 consistent with the backpointer.

1 65. The system according to claim 56, wherein the
2 replicas include replicas of a first type and of a
3 second type wherein locations of replicas of the first
4 type are registered in a parent directory for a file.

1 66. The system according to claim 56, wherein in
2 response to a user accessing a file at a node, a
3 replica of the file is formed at the node.

1 67. The method according to claim 56, wherein a
2 replica is deleted by marking the replica as invalid.

1 68. The method according to claim 67, wherein said
2 marking the replica as invalid comprises removing the
3 backpointer for the replica.

1 69. The method according to claim 67, further
2 comprising periodically removing replicas marked as
3 invalid.